University of Zurich[UZH]

Zurich Open Repository and Archive
University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2016

How institutions shaped the last major evolutionary transition to large-scale human societies

Powers, Simon T ; van Schaik, Carel P ; Lehmann, Laurent

Abstract: What drove the transition from small-scale human societies centred on kinship and personal exchange, to large-scale societies comprising cooperation and division of labour among untold numbers of unrelated individuals? We propose that the unique human capacity to negotiate institutional rules that coordinate social actions was a key driver of this transition. By creating institutions, humans have been able to move from the default 'Hobbesian' rules of the 'game of life', determined by physical/environmental constraints, into self-created rules of social organization where cooperation can be individually advantageous even in large groups of unrelated individuals. Examples include rules of food sharing in hunter-gatherers, rules for the usage of irrigation systems in agriculturalists, property rights and systems for sharing reputation between mediaeval traders. Successful institutions create rules of interaction that are self-enforcing, providing direct benefits both to individuals that follow them, and to individuals that sanction rule breakers. Forming institutions requires shared intentionality, language and other cognitive abilities largely absent in other primates. We explain how cooperative breeding likely selected for these abilities early in the Homo lineage. This allowed anatomically modern humans to create institutions that transformed the self-reliance of our primate ancestors into the division of labour of large-scale human social organization.

DOI: https://doi.org/10.1098/rstb.2015.0098

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-120336
Journal Article
Accepted Version

Originally published at:
Powers, Simon T; van Schaik, Carel P; Lehmann, Laurent (2016). How institutions shaped the last major evolutionary transition to large-scale human societies. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 371(1687):20150098.
DOI: https://doi.org/10.1098/rstb.2015.0098

How institutions shaped the last major evolutionary transition to large-scale human societies

Simon T. Powers*
Carel P. van Schaik†
Laurent Lehmann‡

September 23, 2015

Abstract

What drove the transition from small-scale human societies centred on kinship and personal exchange, to large-scale societies comprising cooperation and division of labour among untold numbers of unrelated individuals? We propose that the unique human capacity to negotiate institutional rules that coordinate social actions was a key driver of this transition. By creating institutions, humans have been able to move from the default "Hobbesian" rules of the "game of life", determined by physical/environmental constraints, into self-created rules of social organisation where cooperation can be individually advantageous even in large groups of unrelated individuals. Examples include rules of food sharing in hunter-gatherers, rules for the usage of irrigation systems in agriculturalists, property rights, and systems for sharing reputation between medieval traders. Successful institutions create rules of interaction that are self-enforcing, providing direct benefits both to individuals that follow them, and to individuals that sanction rule breakers. Forming institutions requires shared intentionality, language, and other cognitive abilities largely absent in other primates. We explain how cooperative breeding likely selected for these abilities early in the *Homo* lineage. This

*Corresponding author. Address: Department of Ecology & Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland. Tel: +41 21 692 4267. Email: Simon.Powers@unil.ch

†Address: Anthropological Institute & Museum, University of Zürich, Zürich, Switzerland. Email: vschaik@aim.uzh.ch

‡Address: Department of Ecology & Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland. Tel: +41 21 692 4183. Email: Laurent.Lehmann@unil.ch

allowed anatomically modern humans to create institutions that transformed the self-reliance of our primate ancestors into the division of labour of large-scale human social organisation.

Keywords: cooperation; institutions; division of labour; human evolution; trade; punishment

Introduction

Life on earth has undergone a series of major evolutionary transitions in which individuals at a lower level of biological organisation came together to form higher-level units [1]. Examples include replicating molecules coming together to form protocells, single-celled individuals evolving into multicellular organisms, and solitary insects transitioning into eusocial colonies. The final transition proposed by Maynard Smith & Szathmáry [1] is the origin of human societies. Yet while the other major evolutionary transitions are starting to become well understood [2, 3], there is a lack of a cohesive theory that can explain the transition from primate social organisation based on kinship and personal exchange to human societies with large-scale impersonal exchange and division of labour between unrelated individuals.

Human societies do indeed largely meet the criteria for a major evolutionary transition [3]. For example, just as epigenetic inheritance (a novel inheritance mechanism) allows the cells in a multicellular organism to differentiate and profit from a division of labour, so language (a novel cultural inheritance mechanism) allows human individuals to coordinate and specialise in different tasks, and so also to profit from a division of labour. Similarly, while by most measures a multicellular organism is more complex than a single cell, so human chiefdoms are more complex than hunter-gatherer bands in terms of the number of hierarchical levels of organisation [4]. And just as multicellular organisms with division of labour and sterile somatic cells gradually evolved from single-celled ancestors, so **the phylogenies of language vocabulary trees** point to states evolving gradually from chiefdoms, which in turn evolved gradually from hunter-gatherer macro-bands and tribes [4].

We propose to subdivide the major transition to large-scale human societies into four distinct, smaller transitions (Fig. 1). (1) The origin of the human hunter-gatherer niche, characterised by large but hard to acquire food packages, allomaternal care, and egalitarian social structure. (2) The origin of language, a novel unlimited inheritance system that

strongly facilitates cumulative cultural evolution and negotiation between individuals. (3) The Neolithic revolution, which involved the shift to agricultural and sedentary populations with hierarchical social organisation. (4) The origin of states, where interactions regularly occur between non-kin who may never meet again.

We will assume that the first transition, from a largely vegetarian primate living in fission-fusion societies in woodland landscapes, to a savannah-living partly carnivorous cooperative hunter type of living, was made possible by changes in social organisation not unlike those seen in other lineages that ended up adopting a combination of cooperative breeding and hunting (e.g. [5]). Our focus here, then, is on explaining the transitions in social organisation subsequent to the emergence of language. Current estimates place the origin of modern-like language at either less than 100 kya or at around 500 kya, with the older date being the most plausible [6].

From an economic point of view, **the major transition is from an initial state of autarky in which group members do not typically exchange resources with each other**, to one of catallaxy where there is extreme division of labour and hence interdependence between group members. **In non-human primate social systems, each individual produces itself most of the resources and technology it needs to survive and reproduce. By contrast, while hunter-gatherer individuals can typically still each produce their own technology, they are reliant on the sharing of food with other individuals to survive. Finally, in large-scale human societies individuals rely on trade with non-kin for nearly all of their vital resources, and an individual will not ~~always~~ itself possesses~~s~~ the entire knowledge necessary to produce any single piece of technology.**

Here we develop the hypothesis that the human capacity to form *institutions* was a key driver of the transition to large-scale societies (and may indeed be necessary for their formation). Institutions (*sensu* [7]) are human-devised mechanisms for generating the rules of social interactions . Through communication and negotiation, humans can transform the rules of their "game of life". The game of life depends on two kinds of constraints. The first kind consists of exogenous biotic and abiotic factors that cannot be changed by individuals at the time they are interacting [7, 8]. These factors include the laws of physics and the current environment, which comprises for example the current total resource endowment and the individuals' state of technology. The second type of constraints are behavioural in nature and

so can potentially be modified by the individuals themselves [7, 8]. This includes restraining or expanding behavioural options. By creating institutions, individuals can change the rules aspect of their social interactions, thereby increasing some possibilities without foreclosing others, and potentially tipping the balance from a situation where defection is individually advantageous into one where it pays to cooperate [9, 10, 8, 11, 12]. As we will discuss below, theoretical work in economics has formally demonstrated conditions under which this can occur even in arbitrarily large groups of unrelated individuals where participants meet very infrequently [10, 11, 12].

Humans can create these institutions because they possess various cognitive features that are lacking in other primates, and that are necessary to devise and enforce institutional rules. These include shared intentionality, strong inhibitory control, and a willingness to seek out mutual opportunities. We explain below how these skills evolved as a result of the adoption of cooperative breeding early in the *Homo* lineage. Once in place, they could then be co-opted for institution formation.

In the remainder of this paper, we first define the term *institution* more precisely, before delineating their costs and benefits, and discussing the cognitive prerequisites necessary for their evolution. We then discuss how the institutional-path hypothesis can explain the key steps of social evolution from hunter-gatherers with language to large-scale states.

Institutions

What is an institution?

In general the outcome(s) of an individual's behaviour, in terms of its fitness consequences and/or material rewards, depends upon the behaviour of other individuals as well as on exogenous biotic and abiotic factors. In game theory [13, 14], a *game form* defines the behavioural options – the "strategies" – available to each individual, and the relationship between strategies and outcomes. The game form thus specifies the rules of social interactions or, in other words, the "rules of the game", which is usually and casually referred to in evolutionary biology as a game. More particularly, in game theory, a game consists of a game form and the preferences of individuals over alternative outcomes, and thus determines

equilibrium strategies [13, 14]. In evolutionary biology, strategies are often (but not always) assumed to be genetically or culturally inherited, in which case it is directly the evolutionary process and not the preferences that determine equilibrium strategies.

When individuals can communicate with each other, and when the strategies consist of messages, a game form is often called a mechanism [7]. We follow Hurwicz [7] in considering that an institution is a mechanism whose outcome is a game form. The hallmark of an institution is a sequence of at least two sets of social interactions:

1. Active genesis of institutional rules through communication and bargaining by the individuals in a group (or subset thereof).

2. Economic interactions whose outcomes are material, and which are affected by the institutional rules.

An institution thus consists of a political game form, which determines the rules of the subsequent economic interactions (Fig. 2). These two types of interaction are likely to take place on very different timescales. In particular, the political game form is likely to be played much less frequently than the economic game form. For example, the economic game form will likely be played many times in a single generation, while the institutional rules may only change once every several generations. The rules of the political game will also themselves be set by rules generated by another game form, referred to as a "constitutional" game form by Ostrom [15, p. 59]. This constitutional game form will in turn be played even less frequently. Finally, the rules of the constitutional game form will themselves be set by a "meta-constitutional" game form, but this series of rule-generating game forms eventually begins with the unchangeable rules of the biophysical world, and terminates with the economic game form that generates material payoffs [15]. Because our focus is on the distinction between generating rules versus playing the economic game form, for simplicity we consider only one political game form and one economic game form.

A more all-encompassing and formal definition of institutions than that given above exists (most notably the one by Hurwicz [7, p. 128]), but for our purposes it is enough to see an institution as a mechanism involving communication whose outcomes are rules of interactions. Non-linguistic animals are probably unable to produce institutions involving many individuals, even though they play economic games, because they cannot autonomously generate institutional rules through communication.

The assemblies in modern hunter-gatherers that discuss resource allocation rules or what would be adequate norms of behaviour provide a good example of an institution as defined above. We stress that the institution comprises the negotiation process as well as the resulting norms or rules of behaviour [7, p. 128]. This is in contrast to the cultural evolution literature, which equates institutions with equilibrium norms of behaviour in an economic game form [16], rather than with a political game form that generates rules for the economic game form.

The benefits of institutions

The formation of institutional rules can transform the "Hobbesian" rules (or default rules) of the game of life into different rules that lead to more cooperative outcomes, but why is this? Since interactions are localised, it is important to realise that social life in hominins largely consists of a repetition of interactions that involve coordination or cooperation problems. For repeated interactions, the fundamental folk theorem of game theory [13, 14, 17] tells us that cooperation can ultimately be sustained in an equilibrium by conditional strategies that respond to players' past actions (reciprocity).

Specifically, in an indefinitely long sequence of interactions where individuals value future payoffs and cannot completely hide their actions, any strategy that guarantees a payoff at least as great as the minimax payoff in the underlying stage game can be an equilibrium [13, 14, 17]. The minimax payoff is the largest payoff that an individual can receive if its opponent tries to minimise the individual's payoff – in the Prisoner's Dilemma it would be the payoff received when the opponent defects. Therefore, if any individual deviates from the equilibrium strategy, then its payoff can be reduced to the minimax payoff by its co-players. Consequently, it does not pay an individual to deviate from a strategy that gives more than its minimax payoff. This logic applies even to groups of infinitely large size where an individual does not interact twice with the same partner, provided that there is a way to transmit sufficient information about the past behaviour of partners [10, 18], i.e. reputation. It also applies to interactions where N-players interact simultaneously [19], such as repeated collective action problems.

There are potentially three kinds of issues that can limit the application of the folk theorem, each of which can be addressed by institutions. The first potential problem stems

from the fact that there are infinitely many cooperative equilibria [17, 20]. Many of these equilibria will give payoffs that are hardly any better than the minimax payoff, while others will result in much greater payoffs. If individuals act independently, then they have no means to guarantee that they will coordinate on an equilibrium that gives high individual payoffs, and are likely to settle on the "default equilibrium" determined by the default Hobbesian rules of interactions (Fig. 3). Institutions can resolve this problem, because they provide a means for individuals to amalgamate dispersed information about resources and wants, and hence coordinate their actions to reach an equilibrium that gives higher payoffs than the default equilibrium (Fig. 3). By devising rules of interactions they settle on an equilibrium, transforming the social contract (in the sense of [17, 20]) from one that gives only the payoff of the Hobbesian equilibrium, to one where the benefits of cooperation are achieved.

The second issue is that individuals need to value future payoffs, and the game needs to be indefinitely repeated. Institutional rules can help to make these conditions hold. For example, Casari [21] describes the development of institutional rules to govern the use of common agriculture land in the Italian Alps, between 1200-1800 AD. The rules which most villages ended up adopting tied families and their future descendants into the group, by requiring that the sale or purchase of rights to use the communal land was subject to a majority vote amongst the other villagers. This ensured that individuals would then care about their future payoffs, and that there was no simple way to end the game.

The third issue is that individuals need to have sufficient information about the past behaviour of other individuals, a problem which becomes all the more pressing as group size increases. Institutional rules can help to alleviate these problems by facilitating the spread of information between group members. For example, extant groups managing common-pool resources from irrigation systems to shared grazing lands make agreements to appoint individuals to act as monitors, and regularly hold assemblies of all group members to share information [9, 21]. Institutional rules that resolve social dilemmas also typically centralised repositories for storing information about the reputation of group members, which was common for merchants in medieval Europe [10, 12]. The right institutional rules, then, can create an environment in which the Folk Theorem can apply [22].

Institutional rules are typically not imposed externally, but are the result of a political game form. Experiments have repeatedly found that individuals placed in social dilemmas and allowed to communicate achieve better outcomes than if they are not allowed to com-

municate [23]. Those using communication both to agree on a joint investment strategy and choose their own sanctioning system achieve results close to the group's optimum ([23], see also [24]). Field studies have illustrated how institutional rules, designed by resource users themselves, allow for the self-organised management of irrigation [25, 26, 27], fishing, and harvesting systems [9]. For example, in the Spanish *huerta* irrigation systems, institutional rules specify how much water each user may take at a given time, how responsibilities for maintenance of the system are shared, and what the sanctions are for individuals that break the rules. These rules are not imposed externally but are created by assemblies of the irrigators themselves, and indeed have been for a thousand years [9].

Critically, and contrary to cultural group selection arguments (e.g. [28, 29, 30]), institutional rules in these situations create a game form in which monitoring and sanctioning are not altruistic (*sensu* evolutionary biology [31]). Rather, field studies have demonstrated that successful institutional rules create conditions that provide *direct benefits* (*sensu* evolutionary biology [31]) to individuals that actively monitor and enforce them [10, 12, 9, 22, 32]. As such, in contrast to altruistic punishment [33, 29], they do not require high genetic or cultural relatedness between group members.

For example, Ostrom describes how extant small-scale societies incentive group members to monitor each other, by allowing individuals that discover a cheater to keep a proportion of the fine levied on that cheater [9]. And as an example in larger-scale societies, in medieval Europe the Law Merchant system of institutional rules was developed, where individuals could pay a cost to register non-cooperative acts by their partner with a judge. They could also pay a cost to query the system to see whether their trading partner had any disputes registered against them before transacting [10, 12]. Judges could impose a fine on cheaters, but had no means to force individuals to pay this fine. Nevertheless, if a fine was imposed then it was in the trader's own interests to pay it in order to maintain a good reputation with the Law Merchant, and so be able to reap the benefits of cooperation with other individuals in future. Consequently, this system of sanctioning was self-enforcing, even though traders could not be compelled to pay a fine, and had to finance the Law Merchant system themselves [10].

The costs of institutions

Creating self-enforcing institutional rules is a costly process. First, there are costs to setting up a self-enforcing system of monitoring and sanctioning, such as paying judges in the Law Merchant system. Second, time and energy must be spent on negotiating the rules. While this can be done in face-to-face discussions after sunset in hunter-gatherer groups, negotiation becomes much more costly as group size increases. Indeed, it cannot be overstated how difficult it is to agree on something in a group. Arrow's impossibility theorem [34] says that there is no satisfactory way of making social decisions once individuals have sufficiently different preferences. As a result, institutional arrangements that need a high level of consensus betwee/n group members may be inherently unstable whenever individuals' endowments and allegiances shift over time, or when there is a turnover of players. Finally, some individuals may exert disproportional influence in the political interactions, driving the creation of rules that favour themselves at the expense of others. The cooperation and coordination achieved under the institutional rules needs to provide sufficient benefits to offset all of these costs, and thus improve on the payoff from the Hobbesian equilibrium (Fig. 3). Nevertheless, the fact that we see cooperation-promoting institutions in the real world implies that this condition can in principle be met.

The uniqueness of institutions in humans

We emphasise here the uniquely human genesis of institutional rules: the explicit and coordinated construction of group-wide rules that regulate social interactions and that are enforced by other group members. This contrasts in a fundamental way to the usual mechanisms for social interactions considered in evolutionary biology. Other organisms can indeed condition their behaviour on the actions of other individuals (e.g. reciprocity), and they can modify their environment over time (niche construction, [35]). Other animals also perform social learning, imitating traditions of other group members such as bird songs or techniques to open nuts. But crucially, we are aware of no other species that over one individual's lifetime can construct arbitrary rules to regulate social activity, and then enforce these rules by coordinated sanctioning (see also [36] for a similar argument about the uniqueness of human culture).

For example, consider the institutional rules of marriage, the details of which are par-

ticular to any one society. At first sight, the reproductive strategy of monogamy adopted by many animals may seem to be the same. But this is not so, because the institutional rules of marriage regulate what counts as marriage, what the necessary preconditions for it to occur are (e.g. the payment of dowries), who may marry who, how a marriage may be terminated, etc. These rules are necessarily recognised and followed by many individuals, and violations are enforced by coordinated sanctioning. In other words, they define what is normative, and they change the rules of the social game by changing the mapping between individual strategies and the corresponding outcomes, i.e. the payoff matrix. By contrast, monogamy in the animal world is simply an individual unilateral reproductive strategy that is not regulated by rules and enforced by societal sanctioning, and so which does not change the rules of the underlying social game.

The cognitive requirements of institutions

It is difficult to see how individuals could play the political game form without certain cognitive faculties that are unique to humans. Institutions involve individuals **bargaining over** rules to structure their social interactions. This means that they first need to be able to foresee alternative social contracts, and then communicate and **negotiate** over them in order to improve over the default Hobbesian rules. This requires at least three types of advanced cognitive features. (1) To devise alternative rules of interactions, individuals need to be able to create virtual worlds. This requires planning, imagination, causal understanding, large working memory, and the ability to anticipate future rewards. (2) To communicate and bargain efficiently over their rules of interactions, individuals need language and a motivation to seek out information and knowledge, have shared intentionality, and a fully developed theory of mind. (3) To reach consensus, individuals need a strong willingness to seek out mutual opportunities, as well as have strong inhibitory control.

These abilities are only partially present in other primates. Why is this? After all, other primates have **large brains** [37] and relatively well developed cognitive faculties. The answer is that many of the traits require at least some degree of prosocial motivations. **Prosocial motivations are lacking in extant great apes, from which we can infer that they were also lacking in the common ancestor of the great apes and *Homo*. In the next section, we present a hypothesis for why prosocial motivations evolved in our lineage, and hence why our lineage evolved a social cognition that could**

later be co-opted for the formation of institutions.

From primate autarky to human catallaxy

The hunter-gatherer niche before language and institutions

Elements of the hunter-gatherer, or forager, niche were gradually assembled over the past 2 million years or so, but details necessarily remain sketchy. Because language must have affected this niche (see below), we will try to reconstruct what it looked like before language arose, based on comparative evidence and on the archaeological and fossil records.

It is known that by the time good documentation of *Homo erectus* is found, at ca. 1.8 million years ago, the basic elements of hunting and gathering were in place [38]. We can infer aspects of the social system [39], including bonding among males (collective defence against large carnivores and subsequent collective acquisition of meat) and male-female friendships (as found in primates in very large groups, e.g. [40]). Large meat packages inevitably meant wider sharing, including with females and immatures. The latter would have increasing difficulty supporting themselves, given the increasing reliance on technology or endurance running, and thus probably required energy inputs from others. Finally, their large brain size, well above the so-called grey ceiling for hominoids [41], suggests energy inputs for reproducing females. In other words, *Homo erectus* showed many elements of extensive allomaternal care [5], i.e. cooperative breeding.

Comparative studies show that cooperative breeding changes the psychology of primates, and indeed other mammals such as elephants and African wild dogs, when compared to their non-cooperatively-breeding sister taxa [42]. These studies imply that cooperative breeding selects for a high social tolerance and prosocial motivations, leading to a marked increase in socio-cognitive abilities [42, 43]. What is unique in *Homo* is that cooperative breeding and the consequent prosocial psychology were added on top of an already existing large-brained ape-like cognitive system inherited from our earlier hominin (australopithecine) ancestors. This created the potential for a more advanced social cognition than that seen in other cooperatively breeding species. One especially pertinent feature of an ape cognitive system is an ability to understand mental states in other

individuals. In great apes this ability seems to only be used in competitive contexts [42]. But when prosocial motivations co-evolved with cooperative breeding in the *Homo* lineage, this existing ability to grasp mental states could start to be used in cooperative contexts. Ultimately, this culminated in the evolution of shared intentionality [42], i.e. the sharing of psychological states between individuals. Shared intentionality in turn underlies many of the other cognitive prerequisites for institution formation, including the use of language to share information [44].

Hunter-gatherers after the advent of language

The origin of language is a complicated and well-studied area, which we do not address directly here. However, we note that once it evolved, language had two key consequences for hunter-gatherer social evolution. First, language made teaching more effective, which provided more scope for cumulative cultural evolution and hence the development of new technologies [45, 46]. Second, once in place language enabled individuals to negotiate their rules of social interactions; that is, to start create institutions for the first time. These two elements came together to produce greater cooperative division of labour among post-language hunter-gatherers.

With the advent of new complex technologies, such as poison-tipped arrows, nets, and traps, it became possible for individuals to hunt large game in much smaller groups than before. Because hunting large game is inherently unpredictable, having multiple hunting parties within a social unit would provide benefits to wider food sharing as an insurance scheme. Even the best hunters benefit from sharing because this reduces the variance in their own and their offspring's daily calorific intake [47]. The improvements in safety and hunting ability led to the break-up of the old group into smaller subgroups, now called bands or camps. But crucially the reputational effects of language allowed bonds to remain strong among camps of the same community, as shown by frequent moves between them. And increasing returns to scale would provide an advantage to sharing with a larger number of individuals, for example by allowing individuals to overproduce food items that they found easy to obtain and exchanging them for other items. Thus, fewer hunters per party supported by a sharing system would potentially massively reduce variance in an individual's yield [48]. Language made it possible for the sharing networks to become larger and therefore more

stable.

Anthropological studies have shown that individuals in extant hunter-gatherer groups consistently devise and use institutional rules to regulate this kind of food sharing. Indeed, extant hunter-gatherers spend much of their time discussing rules of sharing and gossiping about violations of these [47]. In other words, they negotiate institutional rules and enforce these themselves. Examples includes rules that specify who receives what part of a kill and what quantity [36, 49], with defectors being punished by a variety of means from public ridicule through to ostracism and execution [50]. **While some other primates do practise some degree of food sharing, they do not have non-dyadic systems of food sharing which are regulated and enforced by impartial rules that apply to everyone in the group [48]. This suggests that institutional rules are necessary to regulate such systems of sharing [47], and hence that the supporting institutional rules co-evolved with extended food-sharing in hunter-gatherers after the advent of language.**

Hill [36] gives many more examples of institutional rules in extant hunter-gatherer societies. These include rules concerning access to mating partners within the groups (prohibitions and prescriptions on the basis of age, kin, or ritual membership), polygyny (degree allowed and who may practice it), regulation of violent conflict within and between groups, and rules regulating political power (rules of turn-taking in discussions, and rules governing who will be the leader for different social activities). Institutional rules also affect life history, by specifying who must give resources to juveniles.

During the Upper Paleolithic, we also see the emergence of long-distance trade and division of labour beyond food sharing. Evidence for this includes the remains of materials that had been transported hundreds of kilometres from their origins, and the development of new tools that were specialised in performing specific tasks [51]. Trade would have been strongly favoured by the presence of institutions, because already back then it required a mechanism by which individuals could make faithful promises to invest in labour that would only become useful when the finished product was exchanged. While doing this, the individual would produce less food, which would necessitate the pre-existence of rules for food sharing. Finally, the efficiency of long-distance trade would have been greatly enhanced by an institution for using some type of (pre-numismatic) money. The existence of art and other forms of symbolism suggest that Upper Palaeolithic humans had the cognitive abilities

to do this [51].

An important question concerns how institutional rules were formed in Palaeolithic hunter-gatherers. In other words, what form did the political interactions that generated these rules take? Modern hunter-gatherer groups spend much time discussing institutional rules and violations of these around the camp fire after sunset [50]. Furthermore, observations of these groups show egalitarian political interactions. Individuals typically take turns to give their opinions during group discussions [50]. The role of leaders seems to be to help the group to reach a consensus, rather than to force their own opinion upon others, or to benefit materially. Moreover, ethnographic [50] evidence suggests, and archaeological [52] evidence confirms, that status was not hereditary in these ancient mobile hunter-gatherers. This egalitarianism likely evolved and was maintained by a combination of high degrees of social interdependence in obtaining and defending resources, and lethal weapons that reduced the effects of physical differences in strength between would-be dominants and subordinates [53]. Moreover, it would be difficult for any one individual to monopolise meat from large game. Thus, although each individual should be expected to try to craft institutional rules that benefit itself, the egalitarian social structure would have prevented any one individual from being able to benefit itself too much at the expense of the rest of the group. Consequently, the political game was likely to take the form of a mechanism that aggregated the preferences of all group members [54] without resulting in too much dissent.

The origin of agriculture, large social groups, and hierarchy

The origin of agriculture was likely made possible by many factors [55], including the stabilisation of the climate in the Holocene. However, successful agriculture would have necessitated the expansion of the domain of regulation by group institutions. It would require new property rights, to ensure that an individual could not simply have his plants, animals, land, or stored food taken by others [56]. Agriculture would also require institutional rules to prevent the overexploitation of land and other common pool resources [9, 21]. Groups would also face new social dilemmas brought about by new, shared technology, such as the construction and usage of irrigation systems [57]. The existence of institutions therefore placed humans in a unique position to benefit from agriculture.

It is plausible that institutions aimed at solving these problems co-evolved with the de-

mographic expansion of human groups brought about by agriculture. If cooperation problems were solved, then larger group sizes could potentially benefit individuals through both *economies of scale* (increasing returns in material payoff as a function of population size, [58]) and *economies of scope* (increasing returns in material payoff due to variety, not size, [59]). The logic of this has been demonstrated in a formal model of the co-evolution of demography with institutions to regulate irrigation [60]. The results were that groups with institutional rules that successfully solved collective action problems grew to a larger size, and spread their institutional rules to other groups through excess migration.

However, as humans shifted to intensive modes of subsistence the political game form itself started to change [61]. With the advent of storage technologies, it became possible for some individuals to start to build up a surplus of resources and form patrilineal clans for their effective defence. Permanent agriculture, especially irrigation systems, would have tied individuals to their group, making it hard to escape a despotic leader. The result was that agriculture triggered a shift from egalitarianism back to despotic social structure. Despotic leaders that commanded surpluses of resources would then be able to influence institutions for their own good at the expense of other group members, in a way that they could not have done previously in an egalitarian structure. For example, leaders could dominate the political game form and skew the rules in their favour by enforcing (with coalitional support) what proportion of surplus resources from irrigation farming they could keep for themselves rather than share with the rest of the group [62, 63]. Consequently, the shift to intensive food production heralded a transition to coercive and non-egalitarian institutions, or so-called extractive institutions [64].

The origin of states and large-scale markets

Agriculture ultimately led to the emergence of multiple levels of hierarchical organisation – states. In a state the individuals just below the leader in the hierarchy each specialise in just a subset of the tasks of the ruler [65]. The creation of specialised authority roles represents a new institution, i.e. a new political game form that determines the rules of the economic interactions of commoners.

The archaeological evidence shows that the first states arose by the aggregation of previously independent groups, rather than by one group simply expanding in size and displacing

its neighbours [66]. At least two types of between-group interactions are implicated in driving this aggregation: warfare and trade [66]. The role of warfare is quite intuitive: aggregation can happen by one polity forcing another to become subordinate to it. The role of trade is often seen as creating ties between chiefs, through the procurement and exchange of prestige goods (e.g. [67]; see [68] for a review). However, there is also evidence that staple goods were traded over long distances during the Neolithic [69]. Indeed, institutional rules regulated trade during the Neolithic, and prices even followed market rules of supply and demand [68]. Trade would be most reliable, and have lower transaction costs [70], with those others that were playing by the same institutional rules. Institutions could therefore provide a pressure for groups to aggregate into a larger polity in order to reap the gains of economies of scale and scope from trade.

In modern states, division of labour is so pronounced that individuals are critically dependent on others outside their family and close friendship circles for food, as well as for protection from the myriads of hazards encountered during their lifespan. These vital elements for reproduction and survival are often supplied by decentralised competitive markets. This arrangement results in a "mutual dependence among strangers" [71] where there is a remarkable level of trust among interacting individuals, which appears to be as uniquely human as language.

The central problem behind the functioning of any market, and more generally any large-scale society, is that no one has complete information [72]. The rewards of competitive exchange thus seem unachievable without institutional rules that guarantee at least secure property rights, and that enforce the various structural characteristics of information flow in markets [17]. Furthermore, not all markets can be efficient because resource allocations made at the individual level often result in externalities, i.e. the effects of an individual's action can spill over into the environment and negatively affect other individuals, creating a tragedy of the commons situation [73]. This forces groups to design institutional rules that regulate any spillover on the environment, such as the medieval Law Merchant that facilitated trade between strangers by spreading information about their past behaviour [10]. Indeed, the quality of institutional rules has been proposed as being the single main determinant of whether modern nations will fail or succeed. Acemoglu & Robinson [64] distinguish between institutional rules that are *inclusive*, meaning that they provide incentives to individuals that reward them for their productivity, in contrast to *extractive* rules that reward only a

few individuals and that fail to adequately protect property rights.

Discussion

Institutions, the individually devised communication processes that produce rules to structure social interactions, are evolved (extended) phenotypes that fundamentally rely on language. They are key determinants of, and may be necessary to explain, the last major evolutionary transition.

In economics, institutions are often thought of as being formed by individuals with unbounded levels of cognition; that is, individuals following the neoclassical rational choice assumptions [70, chapter 3]. But preferences for forming institutions and institutional rules can also evolve by processes of cultural evolution [60]. Consequently, institutions can be formed by individuals that have only bounded rationality **(like probably any hominoid)**, as long as institution formation increases an individual's own payoff, or their inclusive payoff or fitness. Our hypothesis for the emergence of large-scale societies relies on self-created and self-enforcing institutional rules, which regardless of the exact level of rationality/cognition of individuals involved, provide direct benefits from cooperation and coordination. Under self-enforcing institutional rules cooperation, and monitoring and sanctioning, are not opposed by within-group selection, and are adaptive at the individual level.

There are at least three alternative hypotheses for the evolutionary origin of large-scale human societies. The first rests on individuals performing biased social learning, especially conformity-biased learning, whereby they tend to imitate the most frequent behaviours within their group. This creates high cultural relatedness within groups, and thus enables cultural group selection [74]. Although the conditions under which this really works remain unclear [75, 76, 77], it implicitly assumes a very low level of rationality, because individuals are assumed to be unable to compute when it would actually be advantageous to express shirking behaviours [78]. Consequently, in contrast to the institutional-path hypothesis, the biased-social-learning hypothesis entails that cooperation is often maladaptive at the individual level. Yet although experiments show that human infants develop a propensity for unconditional helping by the age of two, by the age of three they start to become influenced by the past be-

haviour of their partners [79]. In other words, as they develop children do start to take account of expected benefit when deciding whether to cooperate. Such individually-beneficial cooperation is expected under the institutional-path hypothesis. Interestingly, this is the same age at which children start to normatively enforce the rules of artificial games in the laboratory [80].

Moreover, because the cultural evolution literature has essentially ignored the possibility of humans playing a political game, it has implicitly assumed that the rules of the economic game form cannot be changed by a process operating *within* groups. As such, the cultural evolution literature has concluded that rules can only change by a slow and external process of between-group competition, rather than being variable in the short term through internal negotiation. But this conclusion does not fit will with the large brains and advanced planning and negotiation skills of our species.

The second hypothesis presupposes the formation of coercive hierarchy, which results from strong asymmetries in physical strength or power within groups (Fig. 1). Coalitions of powerful individuals (elites) are able to coerce others when surpluses, as produced by agriculture, are large enough to be exploited. They may increase this coercion as groups expand in size through monopolising resources gained through conflict with other groups [81, 82]. This ultimately results in elites creating coercive states through conquest [83]. Under this hypothesis, individuals may still behave in their self-interest when coerced, but the social equilibrium will be far from a group-efficient outcome.

The third hypothesis (the "interdependence hypothesis") is based on the idea that cooperation in early humans was mutualistic, with individuals becoming highly dependant on each other through the scavenging of the carcasses of large game, which later extended into cooperative hunting [84]. This required the development of shared intentionality, and then other advanced socio-cognitive features such as language, in order to ensure successful coordination in high risk Stag-Hunt game situations. The high interdependence of individuals, combined with the possibility of partner choice, provided an incentive for individuals to use reputation when deciding whether to cooperate with an individual. However, this kind of cooperation was threatened as group size expanded, party due to the problem of ~~knowing~~ the reputation of other group members. It is hypothesised

that this problem lead to the adoption of group-wide norms and conventions, and symbolic markers as proxies for reputation [84].

There are clearly strong connections between all these hypotheses, and several elements of them are not mutually exclusive. Both the biased-cultural-transmission and the institutional-path hypotheses rely fundamentally on cultural evolution, and thus involve social learning. The main difference is the conception of rationality that individuals are endowed with. Under the institutional-path hypothesis individuals are assumed to have high levels of cognition and rationality (see section "The cognitive requirements of institutions"), enough at least to respond adaptively to their social environment and reinforce individually beneficial actions under most circumstances. But it does not at all require conformity- or prestige-biased transmission. While conformity is surely important in humans and other primates, we also know that humans are flexible with their investment in cooperation depending upon the context [78, 85, 86], and that there is strong within-culture variation in the social learning strategies that individuals employ [87]. The institutional-path hypothesis better fits with these findings, by not requiring within-group homogeneity of behaviour or preferences.

Having institutions also does not exclude hierarchy and dominance. If hierarchical command is an efficient mode to solve economic problems as group size increases [88], then the voluntary creation of hierarchy and leadership is exactly what we expect to see in the long-run under the institutional path hypothesis. The political game form can then subsequently change into one of dominance, where the new leaders take advantage of the costs of resisting or dispersing to create institutional rules that benefit themselves at the expense of others [62, 63], paving the way to extractive institutions. However, coercive hierarchy seems to be inherently unstable [89] and costly to maintain, given the possibility for subordinates to form coalitions. Moreover, extant small-scale societies demonstrate that egalitarian institutions can resolve social dilemmas in irrigation and other agricultural problems, and often do so more effectively than coercive institutions [9].

Finally, both the institutional-path and interdependence hypotheses agree that human cooperation was driven by direct benefits in small groups. But the interdependence hypothesis argues that the mechanisms supporting this must have broken down in large groups, leading to the use of conformity, group-wide norms and conventions, and symbolic markers as proxies for reputation. How-

ever, this hypothesis does not provide an account of how particular group-wide norms and conventions would be adopted. By contrast, under the institutional-path hypothesis institutional rules continue to provide direct benefits to cooperating even in large groups, and are created by a political game form.

We conclude that the key puzzle about large-scale human societies is not how to explain the existence of altruistic cooperation that is costly and fitness reducing over an individual's lifetime, as has been widely suggested [90]. Instead, the puzzle lies in understanding how the institutional rules that provide lifetime direct benefits to cooperation and coordination are generated and sustained over both short and long times scales. From a theoretical perspective, there is a need for further modelling work on the evolution of institutional rules. From an empirical perspective, future work should investigate further how the cognitive prerequisites for creating institutions evolved, and what the exact level of rationality required is. It should also examine the role of the co-evolution of trade and warfare with institutions, and the concomitant rise of large-scale societies.

References

[1] Maynard Smith, J. & Szathmáry, E., 1995 *Major Transitions in Evolution.* Oxford: W. H. Freeman/Spektrum.

[2] Bourke, A. F. G., 2011 *Principles of Social Evolution.* Oxford Series in Ecology and Evolution. Oxford: Oxford University Press.

[3] Szathmáry, E., 2015 Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences* **112**, 10104–10111. (doi:10.1073/pnas.1421398112).

[4] Currie, T. E., Greenhill, S. J., Gray, R. D., Hasegawa, T. & Mace, R., 2010 Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature* **467**, 801–804. (doi:10.1038/nature09461).

[5] van Schaik, C. P. & Burkart, J. M., 2010 Mind the gap: Cooperative breeding and the evolution of our unique features. In *Mind the Gap: Tracing the Origins of Human Universals* (eds. P. M. Kappeler & J. B. Silk), pp. 477–496. Berlin Heidelberg: Springer.

[6] Dediu, D. & Levinson, S. C., 2013 On the antiquity of language: The reinterpretation

of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology* **4**. (doi:10.3389/fpsyg.2013.00397).

[7] Hurwicz, L., 1996 Institutions as families of game forms. *The Japanese Economic Review* **47**, 113–132. (doi:10.1111/j.1468-5876.1996.tb00038.x).

[8] Gardner, R. & Ostrom, E., 1991 Rules and games. *Public Choice* **70**, 121–149.

[9] Ostrom, E., 1990 *Governing the Commons : The Evolution of Institutions for Collective Action*. Cambridge University Press.

[10] Milgrom, P. R., North, D. C. & Weingast, B. R., 1990 The role of institutions in the revival of trade: The Law Merchant, private judges, and the Champagne Fairs. *Economics & Politics* **2**, 1–23. (doi:10.1111/j.1468-0343.1990.tb00020.x).

[11] Greif, A., Milgrom, P. & Weingast, B. R., 1994 Coordination, commitment, and enforcement: The case of the merchant guild. *Journal of Political Economy* pp. 745–776.

[12] Greif, A., 2006 *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge, UK: Cambridge University Press.

[13] Fudenberg, D. & Tirole, J., 1991 *Game Theory*. Cambridge, MA: MIT Press.

[14] Osborne, J. M. & Rubinstein, A., 1994 *A Course in Game Theory*. Cambridge, MA: MIT Press.

[15] Ostrom, E., 2005 *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press.

[16] Richerson, P. & Henrich, J., 2012 Tribal social instincts and the cultural evolution of institutions to solve collective action problems. *Cliodynamics* **3**, 38–80.

[17] Binmore, K., 2005 *Natural Justice*. Oxford: Oxford University Press.

[18] Kandori, M., 1992 Social norms and community enforcement. *The Review of Economic Studies* **59**, 63–80. (doi:10.2307/2297925).

[19] Fudenberg, D. & Maskin, E., 1986 The folk theorem in repeated games with discounting or with incompete information. *Econometrica* **54**, 533–554.

[20] Binmore, K., 2014 Bargaining and fairness. *Proceedings of the National Academy of Sciences* **111**, 10785–10788. (doi:10.1073/pnas.1400819111).

[21] Casari, M., 2007 Emergence of endogenous legal institutions: Property rights and community governance in the italian alps. *The Journal of Economic History* **67**, 191–226. (doi:10.1017/s0022050707000071).

[22] Guala, F., 2012 Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences* **35**, 1–15. (doi: 10.1017/s0140525x11000069).

[23] Ostrom, E., Walker, J. & Gardner, R., 1992 Covenants with and without a sword: Self-governance is possible. *The American Political Science Review* **86**, 404–417. (doi: 10.2307/1964229).

[24] Baldassarri, D. & Grossman, G., 2011 Centralized sanctioning and legitimate authority promote cooperation in humans. *Proceedings of the National Academy of Sciences* **108**, 11023–11027. (doi:10.1073/pnas.1105456108).

[25] Hunt, R. C., 1988 Size and the structure of authority in canal irrigation systems. *Journal of Anthropological Research* **44**, 335–355.

[26] Trawick, P. B., 2001 Successfully governing the commons: Principles of social organization in an Andean irrigation system. *Human Ecology* **29**, 1–25. (doi: 10.1023/a%253a1007199304395).

[27] Janssen, M. A., Bousquet, F., Cardenas, J.-C., Castillo, D. & Worrapimphong, K., 2012 Field experiments on irrigation dilemmas. *Agricultural Systems* **109**, 65–75. (doi: 10.1016/j.agsy.2012.03.004).

[28] Boyd, R. & Richerson, P. J., 1992 Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology* **13**, 171–195. (doi: 10.1016/0162-3095(92)90032-y).

[29] Boyd, R., Gintis, H., Bowles, S. & Richerson, P. J., 2003 The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences* **100**, 3531–3535.

[30] Bowles, S., Choi, J.-K. & Hopfensitz, A., 2003 The co-evolution of individual behaviors and social institutions. *Journal of Theoretical Biology* **223**, 135–147. (doi: 10.1016/s0022-5193(03)00060-2).

[31] West, S. A., El Mouden, C. & Gardner, A., 2011 Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior* **32**, 231–262. (doi:10.1016/j.evolhumbehav.2010.08.001).

[32] Baumard, N., 2010 Has punishment played a role in the evolution of cooperation? A critical review. *Mind & Society* **9**, 171–192. (doi:10.1007/s11299-010-0079-9).

[33] Fehr, E., Fischbacher, U. & Gächter, S., 2002 Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature* **13**, 1–25. (doi:10.1007/s12110-002-1012-7).

[34] Mas-Colell, A., Whinston, M. D. & Green, J. R., 1995 *Microeconomic Theory*. Oxford: Oxford Unversity Press.

[35] Odling-Smee, F. J., Laland, K. N. & Feldman, M. W., 2003 *Niche Construction: The Neglected Process in Evolution*. Monographs in population biology; no. 37. Princeton University Press.

[36] Hill, K., 2009 Animal "culture"? In *The Question of Animal Culture* (eds. K. N. Laland & B. G. Galef), pp. 269–287. Cambridge, MA: Harvard University Press.

[37] Schoenemann, P. T., 2006 Evolution of the size and functional areas of the human brain. *Annual Review of Anthropology* **35**, 379–406. (doi: 10.1146/annurev.anthro.35.081705.123210).

[38] Foley, R. A., 1987 *Another Unique Species: Patterns in Human Evolutionary Ecology*. Harlow, UK: Longman Scientific & Technical.

[39] Van Schaik, C. P. & Willems, E. R. P., 2015 The social system of Homo erectus: Inferences based on extant primates. *Folia Primatologica* **86**, 374+.

[40] Ostner, J., Vigilant, L., Bhagavatula, J., Franz, M. & Schülke, O., 2013 Stable heterosexual associations in a promiscuous primate. *Animal Behaviour* **86**, 623–631. (doi: 10.1016/j.anbehav.2013.07.004).

[41] Isler, K. & van Schaik, C. P., 2012 How our ancestors broke through the gray ceiling. *Current Anthropology* **53**, S453–S465. (doi:10.1086/667623).

[42] Burkart, J. M., Hrdy, S. B. & Van Schaik, C. P., 2009 Cooperative breeding and human cognitive evolution. *Evol. Anthropol.* **18**, 175–186. (doi:10.1002/evan.20222).

[43] Burkart, J. M. & van Schaik, C. P., 2010 Cognitive consequences of cooperative breeding in primates? *Animal Cognition* **13**, 1–19. (doi:10.1007/s10071-009-0263-7).

[44] Tomasello, M. & Carpenter, M., 2007 Shared intentionality. *Developmental Science* **10**, 121–125. (doi:10.1111/j.1467-7687.2007.00573.x).

[45] Pradhan, G. R., Tennie, C. & van Schaik, C. P., 2012 Social organization and the evolution of cumulative technology in apes and hominins. *Journal of Human Evolution* **63**, 180–190. (doi:10.1016/j.jhevol.2012.04.008).

[46] Fogarty, L., Strimling, P. & Laland, K. N., 2011 The evolution of teaching. *Evolution* **65**, 2760–2770. (doi:10.1111/j.1558-5646.2011.01370.x).

[47] Kaplan, H., Gurven, M., Hill, K. & Hurtado, A. M., 2005 The natural history of human food sharing and cooperation: A review and a new multi–individual approach to the negotiation of norms. In *Moral sentiments and Material Interests: The Foundations of Cooperation in Economic Life* (eds. H. Gintis, S. Bowles, R. Boyd & E. Fehr), pp. 75–113. MIT Press.

[48] Jaeggi, A. V. & Gurven, M., 2013 Natural cooperators: Food sharing in humans and other primates. *Evolutionary Anthropology: Issues, News, and Reviews* **22**, 186–195. (doi:10.1002/evan.21364).

[49] Gurven, M., 2004 To give and to give not: The behavioral ecology of human food transfers. *Behavioral and Brain Sciences* **27**, 543–559. (doi:10.1017/s0140525x04000123).

[50] Boehm, C., 1999 *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Harvard University Press.

[51] Ofek, H., 2001 *Second Nature: Economic Origins of Human Evolution*. Cambridge, UK: Cambridge University Press.

[52] Price, T. D., 1995 Social inequality at the origins of agriculture. In *Foundations of Social Inequality* (eds. T. D. Price & G. M. Feinman), pp. 129–151. New York: Plenum Press.

[53] Gintis, H., van Schaik, C. & Boehm, C., 2015 Zoon politikon. *Current Anthropology* **56**, 327–353. (doi:10.1086/681217).

[54] Conradt, L. & List, C., 2009 Group decisions in humans and animals: A survey. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **364**, 719–742. (doi:10.1098/rstb.2008.0276).

[55] Price, T. D. & Bar-Yosef, O., 2011 The origins of agriculture: New data, new ideas. *Current Anthropology* **52**, S163–S174. (doi:10.1086/659964).

[56] Bowles, S. & Choi, J.-K., 2013 Coevolution of farming and private property during the early holocene. *Proceedings of the National Academy of Sciences* **110**, 8830–8835. (doi:10.1073/pnas.1212149110).

[57] Carballo, D. M., Roscoe, P. & Feinman, G. M., 2014 Cooperation and collective action in the cultural evolution of complex societies. *Journal of Archaeological Method and Theory* **21**, 98–133. (doi:10.1007/s10816-012-9147-2).

[58] Pindyck, R. S. & Rubinfeld, D. L., 2001 *Microeconomics*. Upper Saddle River, NJ: Prentice Hall.

[59] Panzar, J. C. & Willig, R. W., 1981 Economies of scope. *American Economic Review* **71**, 268–272.

[60] Powers, S. T. & Lehmann, L., 2013 The co-evolution of social institutions, demography, and large-scale human cooperation. *Ecology Letters* **16**, 1356–1364. (doi:10.1111/ele.12178).

[61] Earle, T., 1997 *How Chiefs Come to Power: The Political Economy in Prehistory*. Stanford, CA: Stanford University Press.

[62] Hooper, P. L., Kaplan, H. S. & Boone, J. L., 2010 A theory of leadership in human cooperative groups. *Journal of Theoretical Biology* **265**, 633–646. (doi:10.1016/j.jtbi.2010.05.034).

[63] Powers, S. T. & Lehmann, L., 2014 An evolutionary model explaining the Neolithic transition from egalitarianism to leadership and despotism. *Proceedings of the Royal Society B: Biological Sciences* **281**, 20141349. (doi:10.1098/rspb.2014.1349).

[64] Acemoglu, D. & Robinson, 2011 *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. New York: Crown Publishers.

738 [65] Spencer, C. S., 2010 Territorial expansion and primary state formation. *Proceedings of*
739 *the National Academy of Sciences* **107**, 7119–7126. (doi:10.1073/pnas.1002470107).

740 [66] Stanish, C. & Levine, A., 2011 War and early state formation in the northern Titicaca
741 Basin, Peru. *Proceedings of the National Academy of Sciences* **108**, 13901–13906. (doi:
742 10.1073/pnas.1110176108).

743 [67] Levine, A., Stanish, C., Williams, P. R., Chávez, C. & Golitko, M., 2013 Trade and
744 early state formation in the Northern Titicaca Basin, Peru. *Latin American Antiquity*
745 pp. 289–308. (doi:10.7183/1045-6635.24.3.289).

746 [68] Oka, R. & Kusimba, C., 2008 The archaeology of trading systems, part 1: To-
747 wards a new trade synthesis. *Journal of Archaeological Research* **16**, 339–395. (doi:
748 10.1007/s10814-008-9023-5).

749 [69] Haldane, C., 1993 Direct evidence for organic cargoes in the Late Bronze Age. *World*
750 *Archaeology* **24**, 348–360. (doi:10.1080/00438243.1993.9980213).

751 [70] North, D. C., 1990 *Institutions, Institutional Change and Economic Performance (Po-*
752 *litical Economy of Institutions and Decisions)*. Cambridge University Press.

753 [71] Seabright, P., 2010 *The Company of Strangers*. Princeton, NJ: Princeton University
754 Press.

755 [72] Hayek, 1945 The use of knowledge in society. *The American Economic Review* **35**,
756 519–530.

757 [73] Hardin, G., 1968 The tragedy of the commons. *Science* **162**, 1243–1248.

758 [74] Richerson, P. J. & Boyd, R., 2005 *Not by Genes Alone: How Culture Transformed*
759 *Human Evolution*. University of Chicago Press.

760 [75] Lehmann, L., Feldman, M. W. & Foster, K. R., 2008 Cultural transmission can in-
761 hibit the evolution of altruistic helping. *The American Naturalist* **172**, 12–24. (doi:
762 10.1086/587851).

763 [76] André & Morin, O., 2011 Questioning the cultural evolution of altruism. *Journal of*
764 *Evolutionary Biology* **24**, 2531–2542. (doi:10.1111/j.1420-9101.2011.02398.x).

765 [77] Molleman, L., Pen, I. & Weissing, F. J., 2013 Effects of conformism on the cultural
766 evolution of social behaviour. *PLoS ONE* **8**. (doi:10.1371/journal.pone.0068153).

[78] Raihani, N. J. & Bshary, R., 2015 Why humans might help strangers. *Frontiers in Behavioral Neuroscience* **9**. (doi:10.3389/fnbeh.2015.00039).

[79] Warneken, F. & Tomasello, M., 2013 The emergence of contingent reciprocity in young children. *Journal of Experimental Child Psychology* **116**, 338–350. (doi: 10.1016/j.jecp.2013.06.002).

[80] Rakoczy, H., Warneken, F. & Tomasello, M., 2008 The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology* **44**, 875–881. (doi:10.1037/0012-1649.44.3.875).

[81] Seabright, P., 2013 The birth of hierarchy. In *Cooperation and Its Evolution* (eds. K. Sterelny, R. Joyce, B. Calcott & B. Fraser), pp. 109–116. Cambridge, MA: MIT Press.

[82] Sterelny, K., 2014 Cooperation, culture, and conflict. *The British Journal for the Philosophy of Science* pp. axu024+. (doi:10.1093/bjps/axu024).

[83] Carneiro, R. L., 1970 A theory of the origin of the state. *Science* **169**, 733–738. (doi: 10.1126/science.169.3947.733).

[84] Tomasello, M., Melis, A. P., Tennie, C., Wyman, E. & Herrmann, E., 2012 Two key steps in the evolution of human cooperation. *Current Anthropology* **53**, 673–692. (doi: 10.1086/668207).

[85] Gurven, M., Hill, K. & Hillard, K., 2002 From forest to reservation: Transitions in food-sharing behavior among the ache of paraguay. *Journal of anthropological research* **58**, 93–120.

[86] Lamba, S. & Mace, R., 2011 Demography and ecology drive variation in cooperation across human populations. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 14426–14430. (doi:10.1073/pnas.1105186108).

[87] Molleman, L., van den Berg, P. & Weissing, F. J., 2014 Consistent individual differences in human social learning strategies. *Nature Communications* **5**. (doi: 10.1038/ncomms4570).

[88] Johnson, G. A., 1982 Organizational structure and scalar stress. In *Theory and Explanation in Archaeology: The Southampton conference* (eds. C. Renfrew, M. Rowlands & B. A. Segraves-Whallon), pp. 389–420. New York: Academic Press.

797 [89] Tainter, J. A., 1988 *The Collapse of Complex Societies*. Cambridge, UK: Cambridge
798 University Press.

799 [90] Fehr, E. & Gachter, S., 2002 Altruistic punishment in humans. *Nature* **415**, 137–140.
800 (doi:10.1038/415137a).

801 [91] Wittfogel, K. A., 1957 *Oriental Despotism: A Comparative Study of Total Power*. New
802 Haven: Yale University Press.

Figure legends

Figure 1: Sub-transitions of the major transition from small- to large-scale human societies, with the major triggers for subsequent changes indicated, as well as crude estimates of the timing of these transitions. Solving the collective action problems inherent in large-scale agriculture may or may not have involved coercion depending on the society (e.g. over management of irrigation [25, 57, 91]). However, the surpluses provided by agriculture eventually led to hierarchal institutions that were prone to collapse and be reformed [89], culminating in the first states ca 4 kya.

Figure 2: An institution is a mechanism of communication whose output is the rules for economic interactions. As such one is in the presence of a political game form, where the preferences of group members for institutional rules are expressed, followed by an economic game form. The political game form could give equal weight to the preferences of all group members, as in egalitarian institutions, or could give more weight to dominant individuals, as happened with the origin of agriculture. The result of this game is the rules (or game form) for the subsequent economic interactions. The economic interactions determine the fitness or material rewards to individuals, and may for example be a variant of a public goods or coordination game, or an exchange economy where goods are traded.

Figure 3: In the absence of institutions, individuals that engage in repeated social interactions are likely to receive only the payoff corresponding to the default or Hobbesian equilibrium of the game of life (see also [8]). However, when group members can communicate and negotiate an agreed coordination mechanism (i.e. create an institution), they can coordinate in the economic game form on an equilibrium that potentially increases their mutual payoff relative to the Hobbesian equilibrium (shaded area, see also [17, 20]). The resulting equilibrium strategies are self-enforcing, in the sense that it is both individually advantageous to follow them when others are doing so, and individually advantageous to sanction group members that do not follow them. Although this figure shows a group of two individuals for illustration, the size of the shaded area and thus the benefit of having institutions actually increases with the number of interacting individuals. This is because the problems of coordinating on an equilibrium without institutions increase with group size [88].

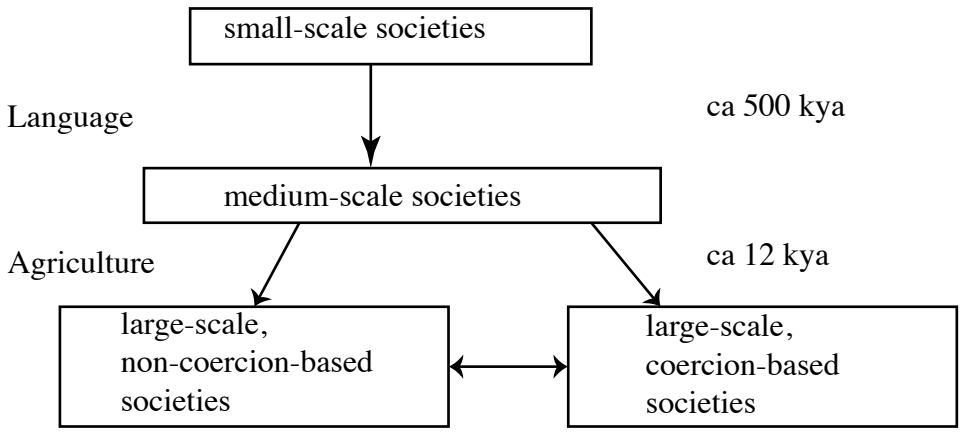
small-scale societies
Language
ca 500 kya
medium-scale societies
Agriculture
ca 12 kya
large-scale,
non-coercion-based
societies
large-scale,
coercion-based
societies

Figure 1

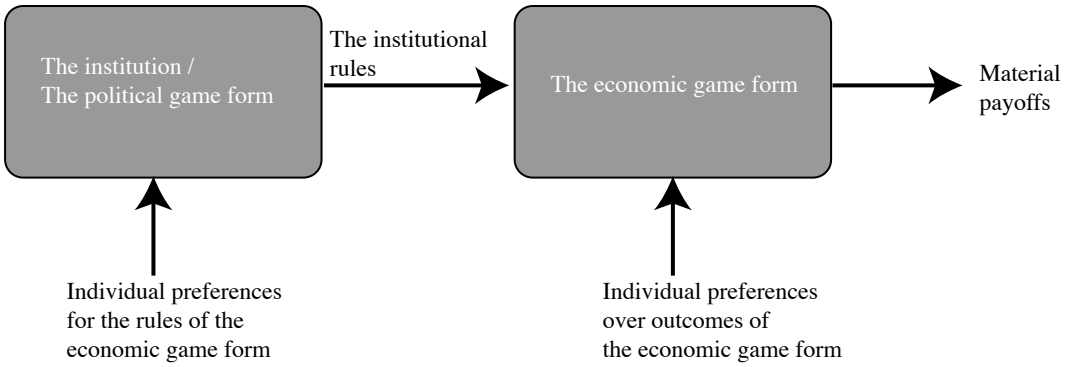
The institution /
The political game form
The institutional
rules
The economic game form
Material
payoffs
Individual preferences
for the rules of the
economic game form
Individual preferences
over outcomes of
the economic game form

Figure 2

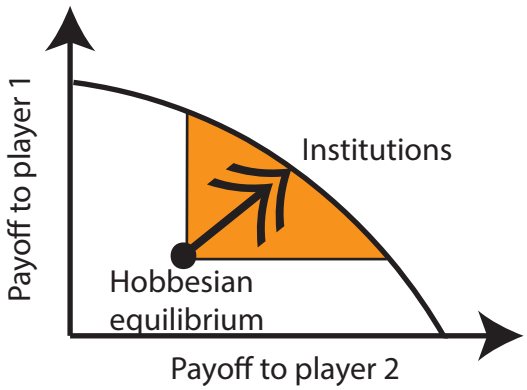
Payoff to player 1
Institutions
Hobbesian
equilibrium
Payoff to player 2

Figure 3